Evaluation of variable numbers of tandem repeat as molecular epidemiological markers of *Mycobacterium tuberculosis* in Japan

Takayuki Wada, Shinji Maeda, Atsushi Hase and Kazuo Kobayashi

1Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, 8-34 Tojo-cho, Tennoji-ku, Osaka 543-0026, Japan
2Molecular Epidemiology and Genetic Identification, Mycobacterium Reference Center, Research Institute of Tuberculosis, 3-1-24 Matsuyama, Kiyose, Tokyo 204-8533, Japan
3Department of Immunology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Using 243 *Mycobacterium tuberculosis* isolates obtained in 2001 in Osaka City, Japan, the discriminatory power of variable numbers of tandem repeats (VNTRs) of 12 standard mycobacterial interspersed repetitive units (MIRUs) was assessed. The biggest cluster defined by MIRU-VNTRs consisted of 57 (23.5%) isolates and they belonged to the Beijing family based on spoligotyping. When additional VNTR loci were included in the MIRU-VNTR analysis, the 57 originally clustered strains were further differentiated by the addition of Queen’s University Belfast (QUB)-VNTRs, but not exact tandem repeat-VNTR. The allelic diversity of additional VNTR loci such as VNTR 3232 (QUB-3232), VNTR 2163a (QUB-11a), VNTR 2163b (QUB-11b) and VNTR 1982 (QUB-18) was high in the 57 strains. When the 243 *M. tuberculosis* isolates were analysed using 16-locus VNTR (the 12 standard MIRUs and the 4 QUB loci) and IS6110 RFLP, the respective Hunter–Gaston discriminatory indexes were 0.9966 and 0.9971. The discrimination power of 16-locus VNTR was equal to that of IS6110 RFLP analysis. If appropriate loci are added to the standard MIRU analysis, VNTR genotyping could be a valuable tool for strain typing and epidemiological research of *M. tuberculosis* in Japan.

INTRODUCTION

Tuberculosis (TB) remains a major public health threat to the world population. The gold standard method for molecular epidemiological analyses for TB is IS6110 RFLP Southern blotting (Cave et al., 1991; Kremer et al., 1999; van Embden et al., 1993). However, it has some limitations, such as the long time for culturing TB bacilli, technical difficulties and the troublesome comparison of data obtained at different laboratories. Therefore, some alternative strategies and methods for epidemiological analysis of TB have been evaluated and reported recently.

Spoligotyping is the most popular PCR-based molecular analytical method in many countries (Kamerbeek et al., 1997). Nevertheless, this method is not informative for *Mycobacterium tuberculosis* Beijing family strains because almost all strains in this family share an identical spoligotyping pattern (Glynn et al., 2002; Kremer et al., 2004). Variable numbers of tandem repeat (VNTR) typing is also PCR-based. It amplifies various loci of minisatellites on a genome and enumerates their units as a molecular marker (Supply et al., 2000). In some reports, the discriminatory capacity of VNTR typing with 12 mycobacterial interspersed repetitive units (MIRUs) is comparable to that of IS6110 RFLP in North America and Europe (Blackwood et al., 2004; Mazars et al., 2001). The Centers for Disease Control and Prevention (CDC) in the USA have adopted it as a standard typing method (Cowen et al., 2005; CDC, 2004).

Although convenient genotyping of *M. tuberculosis* isolates has been achieved, the 12-locus MIRU-VNTR has yielded insufficient discrimination for Beijing strains. Therefore, analyses of various additional VNTR loci for such isolates have been undertaken (Kam et al., 2006; Kremer et al., 2005b; Nikolayevsky et al., 2006; Surikova et al., 2005). Recently, novel standards, 15- or 24-loci MIRU-VNTR methods, were proposed based on analyses of strains from cosmopolitan origins (Supply et al., 2006).
Beijing/W family has been reported to be predominantly spread throughout eastern Asia (Chan et al., 2001; van Soolingen et al., 1995). These strains are also present across former Soviet Union countries (Nikolayevsky et al., 2006) and other areas, like the USA (Bifani et al., 2002). For application of the VNTR method to epidemiological analysis of *M. tuberculosis* in Japan, we performed IS6110 RFLP typing, spoligotyping and 12-locus standard or some additional locus VNTR analyses. These data might be useful to select appropriate VNTR loci for the genotyping of *M. tuberculosis* in Japan.

**METHODS**

**Bacterial strains and DNA isolation.** A total of 243 *M. tuberculosis* strains were collected at the Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, from culture-confirmed pulmonary TB patients who were reported to the public health centre of Osaka, Osaka City, Japan, between January 1 and December 31 2001. All the available isolates collected in the period were used for this study. They were obtained mainly from three hospitals and two medical centres in Osaka. The mean incidence of TB in Japan is 27.9 per 100,000 population, but the incidence is much higher in Osaka City (82.6 per 100,000 population). Mycobacterial genomic DNA was prepared from bacteria grown on Ogawa medium using a combination of phenol/chloroform extraction and mechanical disruption. The DNA concentration was estimated using UV absorbance at 260 nm; then it was adjusted to 20 μg ml⁻¹.

**Molecular typing methods.** IS6110 RFLP typing was performed according to a standardized protocol (van Embden et al., 1993). The band patterns were analysed using dedicated software (Molecular Analyst; Bio-Rad Laboratories). The spoligotyping method (Kamerbeek et al., 1997) was also performed according to a standard protocol (Molhuizen et al., 1998), except for the use of positively charged nylon membranes and DIG-labelled probes (Roche Diagnostic Systems) in place of biotin-labelled probes.

For this study, VNTR typing was performed with Ex Taq and GC PCR buffer 1 (TaKaRa Holdings). The sequences of the primers for amplification of 12 loci of MIRUs (MIRU-02, 10, 16, 20, 23, 24, 26, 27, 31, 39 and 40), 4 loci of exact tandem repeats (ETRs) (ETR-A, B, C and F) and 8 loci of Queen’s University Belfast (QUBs)-VNTR (QUB-11a, 11b, 15, 18, 23, 26, 1895, 3232 and 3336) were selected according to descriptions in other studies (Frothingham & Meeker-O’Connell, 1998; Kremer et al., 2001; Skuce et al., 2002; Supply et al., 2001). The PCR mixture was prepared in a 20 μl volume with 1 × GC PCR buffer 1, 0.005 U Ex Taq, 200 μM each of the four dNTPs, 0.4 μM each primer set and 10 ng template DNA. The PCR conditions were as follows: initial denaturation at 95 °C for 10 min, and then 35 cycles of 95 °C for 1 min, 60 °C for 1 min and 72 °C for 1 min.

**VNTR analysis of amplified DNA fragments.** The sizes of amplified DNA fragments were determined using capillary array electrophoresis analysis equipment (i-chip SV1210; Hitachi Electronics Engineering). Using this equipment, sizes of DNA fragments were calculated digitally in six minutes. Their respective copy numbers were calculated from their size and assigned according to the number of repeats for each locus, and in agreement to published allelic tables (Supply et al., 2001) or as described elsewhere (Frothingham & Meeker-O’Connell, 1998; Kremer et al., 2005b; Skuce et al., 2002; Supply et al., 2001). The accuracy of calculation was determined through comparison with a DNA ladder, as a fragment-size standard, and confirmed through analysis of H37Rv.

**Allelic diversity and discrimination.** The allelic diversity (h) at each VNTR locus was calculated using the equation $h = 1 - \Sigma x_i^2$, where $x_i$ is the frequency of the $i$th allele at the locus, as used in previous works (Kremer et al., 2005b; Sun et al., 2004). The Hunter–Gaston discriminatory index (HGDI) was calculated in accordance with a method, explained in another paper, to evaluate the combination of some VNTR loci (Hunter & Gaston, 1988).

**RESULTS**

**Genotyping analyses of total isolates**

All the 243 isolates were analysed using spoligotyping, MIRU-VNTR and IS6110 RFLP analyses. The numbers of clusters observed when performing IS6110 RFLP and 12-loci MIRU-VNTR were roughly equal (27 in RFLP and 28 in VNTR) (Table 1). However, the respective percentages of clustering isolates by RFLP and VNTR analyses were 30.9 and 70.8 %. The biggest cluster sizes were 7 (2.8 %) for RFLP and 57 (23.5 %) for 12-locus MIRU-VNTR. The respective HGDI values of MIRU-VNTR and IS6110 RFLP were 0.928 and 0.997. The discriminatory power of spoligotyping was lower (0.585), in the population, because it had no discrimination on 154 (63.4 % of strains) Beijing genotype isolates that shared an identical spoligotyping pattern. Moreover, 39 isolates were identified as Beijing-like strains based on the definition in a spoligotyping pattern proposed previously (Kremer et al., 2004; Supplementary Table S1 available with the online journal).

**Analyses of 57 MIRU-identical isolates**

Analyses by spoligotyping and MIRU-VNTR showed that the 57 isolates in the biggest MIRU-VNTR defined cluster belonged to the Beijing family (Supplementary Table S1 available with the online journal) and possessed identical

<table>
<thead>
<tr>
<th>Typing method</th>
<th>Total no. of type patterns</th>
<th>No. of unique types</th>
<th>No. of clusters</th>
<th>No. of clustered isolates (%)</th>
<th>Maximum no. of isolates in a cluster</th>
<th>HGDI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS6110 RFLP</td>
<td>195</td>
<td>168</td>
<td>27</td>
<td>75 (30.9)</td>
<td>7</td>
<td>0.997</td>
</tr>
<tr>
<td>Spoligotyping</td>
<td>38</td>
<td>25</td>
<td>13</td>
<td>218 (89.7)</td>
<td>154</td>
<td>0.585</td>
</tr>
<tr>
<td>12-loci MIRU-VNTR</td>
<td>99</td>
<td>71</td>
<td>28</td>
<td>172 (70.8)</td>
<td>57</td>
<td>0.928</td>
</tr>
</tbody>
</table>

*HGDI was calculated as described by Hunter & Gaston (1988).
MIRU genotypes (223325173533, sequentially in MIRU locus number). *M. tuberculosis* strains with this specific MIRU genotype accounted for 29.5% (57/193) of the Beijing and Beijing-like strains in Osaka. Using these MIRU-identical isolates, we analysed the discriminatory abilities of other VNTR loci. From the $h$ index, high allelic diversity among the 57 strains was observed at VNTR 3232 (QUB-3232), VNTR 2163a (QUB-11a), VNTR 2163b (QUB-11b) and VNTR 1982 (QUB-18) (Table 2). On the other hand, the four ETR loci that were the standard loci of additional VNTR analyses were insufficient to discriminate these strains and offered a low allelic diversity in this cluster (showing an $h$ index of 0 for ETR C and 0.306 for ETR F). Using our PCR conditions, the locus VNTR 3155 (QUB-15) could not be amplified (data not shown).

**Table 2.** The allelic profiles of each additional VNTR locus in 57 MIRU-identical strains

<table>
<thead>
<tr>
<th>VNTR locus</th>
<th>No amplification</th>
<th>Copy no. of repetitive unit(s)</th>
<th>$h$ index*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>VNTR 2165 (ETR A)</td>
<td>1 3 1 52</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td>VNTR 2461 (ETR B)</td>
<td>2 51 3</td>
<td>0.132</td>
<td></td>
</tr>
<tr>
<td>VNTR 577 (ETR C)</td>
<td>57</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>VNTR 3239 (ETR F)</td>
<td>1§ 6 47 3</td>
<td>0.306</td>
<td></td>
</tr>
<tr>
<td>VNTR 2163a (QUB-11a)</td>
<td>1 16 1 5 23 1</td>
<td>0.740</td>
<td></td>
</tr>
<tr>
<td>VNTR 2163b (QUB-11b)</td>
<td>4 5 3 9 13</td>
<td>0.811</td>
<td></td>
</tr>
<tr>
<td>VNTR 1982 (QUB-18)</td>
<td>6 19 1 28 3</td>
<td>0.633</td>
<td></td>
</tr>
<tr>
<td>VNTR 1612 (QUB-23)</td>
<td>1 55 1</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>VNTR 4052 (QUB-26)</td>
<td>2 2 3 41 8 1</td>
<td>0.457</td>
<td></td>
</tr>
<tr>
<td>VNTR 1895 (QUB-1895)</td>
<td>5 52</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td>VNTR 3232 (QUB-3232)</td>
<td>2 2 1 9 6 14 8 15</td>
<td>0.851</td>
<td></td>
</tr>
<tr>
<td>VNTR 3336 (QUB-3336)</td>
<td>1 3 1 42 5 5</td>
<td>0.438</td>
<td></td>
</tr>
</tbody>
</table>

* $h$ index represents the allelic diversity of each locus, calculated as described in Methods.
† The number of strains that had more than 16 copies in the locus are represented all together in this column.
§ The deletion of 3′-terminal 24 bp of a unit, which is present in nearly all isolates, was absent from this isolate.

**Assessment of the combination of additional VNTR loci**

To confirm the combination of four QUB loci (VNTR 3232, VNTR 2163a, VNTR 2163b and VNTR 1982) as an additional VNTR loci set, their genotypes were compared with *IS6110* RFLP analysis in 57 MIRU-identical strains. Among 11 clusters defined by the 4 QUB loci, 7 clusters were completely consistent with clusters identified using *IS6110* RFLP analysis (data not shown). Isolates in each of the remaining four clusters showed nearly identical RFLP patterns (Fig. 1). A one band difference was observed in three clusters (clusters A, B and C) and the isolates in the remaining cluster D showed a two band difference. The discriminatory power of these four additional loci was

---

**Fig. 1.** Differences of *IS6110* RFLP in isolates sharing the same QUB(4)-VNTR genotype. *IS6110* RFLP genotypes were compared for four clusters (A, B, C and D) that represented the same QUB(4)-VNTR types but different *IS6110* RFLP types in 57 MIRU-identical strains. Arrowheads indicate the differences in band pattern(s) of respective clusters.
confirmed by analysing the remaining strains from the collection (186 isolates). The diversity of 4 loci, which offered high allelic diversity in 57 MIRU-identical strains, was also high in the 243 total strains (Table 3). We performed 16-loci VNTR [MIRU(12) + QUB(4)] analyses for the remaining 186 TB isolates to assess the efficiency of the addition of 4 loci to MIRU-VNTR. The VNTR profiles of all 243 isolates are listed in Supplementary Table S1 available with the online journal. The respective percentages of clustering isolates by RFLP and VNTR analyses were 30.9 and 36.6 %. The maximum cluster sizes of these analyses were seven for RFLP and eight for VNTR (Table 1, 4). The HGDI values were similar (0.9971 for RFLP and 0.9966 for VNTR).

**DISCUSSION**

The search for more appropriate genetic markers for the discrimination of *M. tuberculosis* isolates has led to the evaluation and selection of additional VNTR loci designated ETRs, QUBs, Mtub and other loci (Kremer et al., 2005a; Le Fleche et al., 2002; Roring et al., 2002, 2004; Skuce et al., 2002; Smittipat et al., 2005; Warren et al., 2004). For discriminating Beijing strains, a high efficiency of VNTR loci showing hyper-variability, such as some QUBs, has been reported (Kam et al., 2006; Kremer et al., 2005b; Nikolayevsky et al., 2006; Surikova et al., 2005). The proposed optimized MIRU-VNTR typing excludes these loci because of their size allele designation and hyper-variability (Supply et al., 2006). In our data, VNTR 2163a and VNTR 2163b could not be amplified from 13 and 7 isolates (5.3 and 2.9 % of strains), respectively. Moreover, 59 isolates (24.3 %) had copy numbers greater than 16 for VNTR 3232 (Table 3). These results pose a problem for the reliable sizing of amplicons and efficient allele assignment. However, in agreement to the results obtained, VNTR analyses of hyper-variable loci might be useful in some cases to discriminate W-Beijing strains that possess high clonality. Oeleman et al. (2006) used a combination of novel optimized VNTR and spoligotyping, which might be inappropriate for discrimination of strains of Beijing family.

Our data showed that 193 isolates (79.4 % of the strains) were Beijing family strains, suggesting that such strains are prevalent in Japan, as in other East Asian areas (Bifani et al., 2002; Chan et al., 2001; van Soolingen et al., 1995). The discrimination by 12-loci MIRU-VNTR was low in Osaka because of the high prevalence of Beijing strains. This result indicates that, at least, 12-locus MIRU-VNTR is of limited value as a genotyping method of TB in Japan. We selected four QUB loci (VNTR 3232, VNTR 2163a, VNTR 2163b and VNTR 1982) that offered high allelic diversity, within the cluster, to adapt MIRU-VNTR analysis as a promising genotyping method for TB in Japan. The allelic diversity of these loci was consistent with that described in other reports related to analyses of Beijing strains (Kam et al., 2006; Kremer et al., 2005b; Nikolayevsky et al., 2006; Surikova et al., 2005).

**Table 3.** The allelic profiles of four candidates for additional VNTR loci in 243 *M. tuberculosis* strains from Osaka City

<table>
<thead>
<tr>
<th>VNTR locus</th>
<th>No amplification</th>
<th>Copy no. of repetitive unit(s)</th>
<th>h index*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>VNTR 2163a (QUB-11a)</td>
<td>13</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>VNTR 2163b (QUB-11b)</td>
<td>7</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>VNTR 1982 (QUB-18)</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>VNTR 3232 (QUB-3232)</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

* h index represents the allelic diversity of each locus, calculated as described in Methods.
† The number of strains that had more than 16 copies in the locus are represented all together in this column.

**Table 4.** The discriminatory power of 16-loci VNTR [MIRU + QUB(4)] in 243 *M. tuberculosis* strains from Osaka City

<table>
<thead>
<tr>
<th>Typing method</th>
<th>Total no. of type patterns</th>
<th>No. of unique types</th>
<th>No. of clusters</th>
<th>No. of clustered isolates (%)</th>
<th>Maximum no. of isolates in a cluster</th>
<th>HGDI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIRU + QUB(4) VNTR†</td>
<td>188</td>
<td>154</td>
<td>34</td>
<td>89 (36.6)</td>
<td>8</td>
<td>0.9966</td>
</tr>
</tbody>
</table>

*HGDI was calculated as described by Hunter & Gaston (1988).
† QUB(4) VNTR included QUB-11a, QUB-11b, QUB-18 and QUB-3232. According to new nomenclature, these are equivalent to VNTR 2163a, VNTR 2163b, VNTR 1982 and VNTR 3232, respectively.
The QUB(4) loci were highly discriminatory in the 57 MIRU-identical clustered Beijing isolates, the performance was almost equivalent to that of IS6110 RFLP. The clusters A, B and C in Fig. 1 showed only one band difference in IS6110 RFLP. These clusters are therefore likely to represent transmission clusters. These clusters were also consistent with MIRU(12) + QUB(4) defined clusters, suggesting the usefulness of the addition of QUB(4) VNTR typing in defining transmission clusters. In all of the 243 strains, the cluster analysis and the discriminatory power of MIRU(12) + QUB(4) were comparable to that of IS6110 RFLP (Table 4). These results demonstrated that the VNTR method can be used as an alternative for IS6110 RFLP typing in Japan.

Our approach for evaluating an alternative VNTR set, as a genotyping marker system for TB surveillance, was the addition of optional loci to a standard loci set to supplement its low discriminatory power for Beijing family strains. Such a strategy might facilitate both international comparison of epidemiological information and local discrimination of endemic isolates such as strains of the Beijing family.

ACKNOWLEDGEMENTS

The authors are grateful to A. Tamaru, Department of Microbiology, Osaka Prefectural Institute of Public Health, Osaka, Japan, for assistance with DNA electrophoresis using i-chip SV1210 apparatus. This work was supported by grants from the Japanese Ministry of Health, Labour and Welfare (Research on Emerging and Re-emerging Infectious Diseases, Health Sciences Research Grants), Ministry of Education, Culture, Sports, Science and Technology, Ministry of the Environment (Global Environment Research Fund), United States – Japan Cooperative Medical Science Program against Tuberculosis and Leprosy, and Daido Life Welfare Foundation and Ohyama Health Foundation.

REFERENCES


